

Attachment A

Title

A UNIFIED SYSTEM AND METHOD FOR ANIMAL BEHAVIOR
CHARACTERIZATION IN HOME CAGES USING VIDEO ANALYSIS

**A UNIFIED SYSTEM AND METHOD FOR ANIMAL BEHAVIOR
CHARACTERIZATION IN HOME CAGES USING VIDEO ANALYSIS**

Inventors: Xuesheng Bai, Vikrant Kobla, Yiqing Liang, Yi Zhang

5

GOVERNMENT RIGHTS NOTICE

Portions of the material in this specification arose as a result of Government support under grants MH58964, MH58964-02, and DA14889 between Clever Sys., Inc. and The National Institute of Mental Health, National Institute on Drug Abuse,
10 National Institute of Health. The Government has certain rights in this invention.

BACKGROUND OF THE INVENTION

This application is a continuation in part of application No. 09/718,374 filed on November 24 2000, which is now partly allowed.

15

1. Technical Field

The invention relates generally to behavior analysis of animal objects. More particularly, one aspect of the invention is directed to monitoring and characterization of behaviors under specific behavioral paradigm experiments, including home cage behavior paradigms, locomotion or open field paradigm experiment, object
20 recognition paradigm experiments, variety of maze paradigm experiments, water maze paradigm experiments, freezing paradigm experiments for conditioned fear, for an animal, for example, a mouse or a rat, using video analysis from a top view image or side view image, or the integration of both views.

2. Background Art

25

Animals, for example mice or rats, are used extensively as human models in

the research of drug development; genetic functions; toxicology research; understanding and treatment of diseases; and other research applications. Despite the differing lifestyles of humans and animals, for example mice, their extensive genetic and neuroanatomical homologies give rise to a wide variety of behavioral processes that are widely conserved between species. Exploration of these shared brain functions will shed light on fundamental elements of human behavioral regulation. Therefore, many behavioral test experiments have been designed on animals like mice and rats to explore their behaviors. These experiments include, but not limited to, home cage behaviors, open field locomotion experiments, object recognition experiments, a variety of maze experiments, water maze experiments, and freezing experiments for conditioned fear.

Animal's home cage activity patterns are important examination item on the general health list of animals, such as mice and rats. It provides many important indications of whether the animal's health status is normal or abnormal. Home cage behaviors are best observed by videotaping several 24-hour periods in the animal housing facility, and subsequent scoring of the videotape by two independent observers. However, this observation has rarely been done until our inventions came into play, due to the instability in long term human observation, the time consumed, and the huge costs associated with the observation.

As discussed, all these apparatus and experiments use, in many cases, human observation of videotapes of the experiment sessions, resulting in inaccuracy, subjectivity, labor-intensive, and thus expensive experiments. Some automating software provides rudimentary and basic parameters, relying on tracking animal as a point in space, generating experiment results that are inaccurate and can not meet the

demands for advanced features. Besides, each system software module works for only a specific experiment, resulting in potential discrepancy in the results across different systems due to differences in software algorithms used.

5 All the observations of these behavioral experiments use video to record experiment processes and rely on human observations. This introduces the opportunity to utilize the latest technologies development in computer vision, image processing, and digital video processing to automate the processes and achieve better results, high throughput screening, and lower costs. Many of these experiments are conducted with observations performed from top view, that is, observation of the
10 experiments from above the apparatus is used to obtain needed parameters. This also provides an opportunity to unify the approaches to observe and analyze these experiments' results.

SUMMARY OF THE INVENTION

15 There are strong needs for automated systems and software that can automate the measurements of the experiments mentioned above, provide the measurements of meaningful complex behaviors and new revealing parameters that characterize animal behaviors to meet post-genomic era's demands, and obtain consistent results using novel approaches.

20 A revolutionary approach is invented to automatically measure animal's home cage activity patterns. This approach consists of defining a unique set of animal's, such as mice or rats, behavior category. This category includes behaviors like rearing, walking, grooming, eating, drinking, jumping, hanging, etc. Computer systems are designed and implemented that can produce digital video files of animal's

behaviors in a home cage in real time or off-line mode. Software algorithms are developed to automatically understand and analyze the animal's behaviors in those video files. This analysis is based on the premise that the entire animal body, body parts, related color information, and their dynamic motion are taken advantage of in order to provide the measurement of complex behaviors and novel parameters.

In general, the present invention is directed to systems and methods for finding patterns of behaviors and/or activities of an animal using video. The invention includes a system with a video camera connected to a computer in which the computer is configured to automatically provide animal identification, animal motion tracking (for moving animal), animal shape, animal body parts, and posture classification, and behavior identification. Thus, the present invention is capable of automatically monitoring a video image to identify, track and classify the actions of various animals and their movements. The video image may be provided in real time from a camera and/or from a storage location. The invention is particularly useful for monitoring and classifying mice or rats behavior for testing drugs and genetic mutations, but may be used in a number of surveillance or other applications.

In one embodiment the invention includes a system in which an analog/digital video camera and a video record/playback device (e.g., VCR) are coupled to a video digitization/compression unit. The video camera may provide a video image containing an animal to be identified. The video digitization/compression unit is coupled to a computer that is configured to automatically monitor the video image to identify, track and classify the actions of the animal and its movements over time within a sequence of video session image frames. The digitization/compression unit may convert analog video and audio into, for example, MPEG or other formats. The

computer may be, for example, a personal computer, using either a Windows platform or a Unix platform, or a Macintosh computer and compatible platform. The computer is loaded and configured with custom software programs (or equipped with firmware) using, for example, MATLAB or C/C++ programming language, so as to
5 analyze the digitized video for animal identification and segmentation, tracking, and/or behavior/activity characterization. This software may be stored in, for example, a program memory, which may include ROM, RAM, CD ROM and/or a hard drive, etc. In one variation of the invention the software (or firmware) includes a unique background subtraction method which is more simple, efficient, and
10 accurate than those previously known.

In operation, the system receives incoming video images from either the video camera in real time or pre-recorded from the video record/playback unit. If the video is in analog format, then the information is converted from analog to digital format and may be compressed by the video digitization/compression unit. The digital video
15 images are then provided to the computer where various processes are undertaken to identify and segment a predetermined animal from the image. In a preferred embodiment the animal is a mouse or rat in motion with some movement from frame to frame in the video, and is in the foreground of the video images. In any case, the digital images may be processed to identify and segregate a desired (predetermined)
20 animal from the various frames of incoming video. This process may be achieved using, for example, background subtraction, mixture modeling, robust estimation, and/or other processes.

The shape and location of the desired animal is then tracked from one frame or scene to another frame or scene of video images. The body parts of the animal

such as head, mouth, tail, ear, abdomen, lower back, upper back, forelimbs, and hind limbs, are identified by novel approaches through body contour segmentation, contour segment classification, and relaxation labeling. Next, the changes in the shapes, locations, body parts, and/or postures of the animal of interest may be identified, their features extracted, and classified into meaningful categories, for example, vertical positioned side view, horizontal positioned side view, vertical positioned front view, horizontal positioned front view, moving left to right, etc. Then, the shape, location, body parts, and posture categories may be used to characterize the animal's activity into one of a number of pre-defined behaviors. For example, if the animal is a mouse or rat, some pre-defined normal behaviors may include sleeping, eating, drinking, walking, running, etc., and pre-defined abnormal behavior may include spinning vertical, jumping in the same spot, etc. The pre-defined behaviors may be stored in a database in the data memory. The behavior may be characterized using, for example, approaches such as rule-based label analysis, token parsing procedure, and/or Hidden Markov Modeling (HMM). Further, the system may be constructed to characterize the object behavior as new behavior and particular temporal rhythm.

In another embodiment of the invention, there are multiple cameras taking video images of experiment cages that contain animals. There is at least one cage, but as many as the computer computing power allows, say four (4) or sixteen (16) or even more, can be analyzed. Each cage contains at least one animal or multiple animals. The multiple cameras may be taking video from different points of views such as one taking video images from the side of the cage, or one taking video images from the top of the cage. When video images are taken of multiple cages and devices

containing one or multiple animals, and are analyzed for identifying these animals' behaviors, high throughput screening is achieved. When video images taken from different points of views, for example, one from the top view and another from the side view, are combined to identify animal's behaviors, integrated analysis is achieved.

In another preferred embodiment directed toward video analysis of animals such as mice or rats, the system operates as follows. As a preliminary matter, normal postures and behaviors of the animals are defined and may be entered into a Normal Paradigm Parameters, Postures and Behaviors database. In analyzing, in a first instant, incoming video images are received. The system determines if the video images are in analog or digital format and input into a computer. If the video images are in analog format they are digitized and may be compressed, using, for example, an MPEG digitizer/compression unit. Otherwise, the digital video image may be input directly to the computer. Next, a background may be generated or updated from the digital video images and foreground objects detected. Next, the foreground animal features are extracted. Also, body parts such as head, tail, ear, mouth, forelimbs, hind limbs, abdomen, and upper and lower back, are identified. Two different methods are pursuing from this point, depending on different behavior paradigms. In one method, the foreground animal shape is classified into various categories, for example, standing, sitting, etc. Next, the foreground animal posture is compared to the various predefined postures stored in the database, and then identified as a particular posture or a new (unidentified) posture. Then, various groups of postures and body parts are concatenated into a series to make up a foreground animal behavior compared against the sequence of postures, stored in for

example a database in memory, that make up known normal or abnormal behaviors of the animal. The abnormal behaviors are then identified in terms of known abnormal behavior, new behavior and/or daily rhythm. In another method, behavioral processes and events are detected, and behavior parameters are calculated. These behaviors
5 parameters give indications to animal health information related to learning and memory capability, anxiety, and relations to certain diseases.

In one variation of the invention, animal detection is performed through a unique method of background subtraction. First, the incoming digital video signal is split into individual images (frames) in real-time. Then, the system determines if the
10 background image derived from prior incoming video needs to be updated due to changes in the background image or a background image needs to be developed because there was no background image was previously developed. If the background image needs to be generated, then a number of frames of video image, for example 20, will be grouped into a sample of images. Then, the system creates a
15 standard deviation map of the sample of images. Next, the process removes a bounding box area in each frame or image where the variation within the group of images is above a predetermined threshold (i.e., where the object of interest or moving objects are located). Then, the various images within the sample less the bounding box area are averaged. Final background is obtained by averaging 5-10
20 samples. This completes the background generation process. However, often the background image does not remain constant for a great length of time due to various reasons. Thus, the background needs to be dynamically recalculated periodically as above or it can be recalculated by keeping track of the difference image and note any sudden changes. The newly dynamically generated background image is next

subtracted from the current video image(s) to obtain foreground areas that may include the object of interest.

Next, the object identification/detection process is performed. First, regions of interest (ROI) are obtained by identifying areas where the intensity difference generated from the subtraction is greater than a predetermined threshold, which constitute potential foreground object(s) being sought. Classification of these foreground regions of interest will be performed using the sizes of the ROIs, distances among these ROIs, threshold of intensity, and connectedness, to thereby identify the foreground objects. Next, the foreground object identification/detection process may be refined by adaptively learning histograms of foreground ROIs and using edge detection to more accurately identify the desired object(s). Finally, the information identifying the desired foreground object is output. The process may then continue with the tracking and/or behavior characterization step(s).

Development activities have been completed to validate various scientific definitions of mouse behaviors and to create novel digital video processing algorithms for mouse tracking and behavior recognition, which are embodied in a software and hardware system according to the present invention. An automated method for analysis of mouse behavior from digitized 24 hours video has been achieved using the present invention and its digital video analysis method for object identification and segmentation, tracking, and classification. Several different methods and their algorithms, including Background Subtraction, Probabilistic approach with Expectation-Maximization, and Robust Estimation to find parameter values by best fitting a set of data measurements and results proved successful.

The need for sensitive detection of novel phenotypes of genetically manipulated or drug-administered mice demands automation of analyses. Behavioral phenotypes are often best detected when mice are unconstrained by experimenter manipulation. Thus, automation of analysis of behavior in a known environment, for example a home cage, would be a powerful tool for detecting phenotypes resulting from gene manipulations or drug administrations. Automation of analysis would allow quantification of all behaviors as they vary across the daily cycle of activity. Because gene defects causing developmental disorders in humans usually result in changes in the daily rhythm of behavior, analysis of organized patterns of behavior across the day may also be effective in detecting phenotypes in transgenic and targeted mutant mice. The automated system may also be able to detect behaviors that do not normally occur and present the investigator with video clips of such behavior without the investigator having to view an entire day or long period of mouse activity to manually identify the desired behavior.

The systematically developed definition of mouse behavior that is detectable by the automated analysis according to the present invention makes precise and quantitative analysis of the entire mouse behavior repertoire possible for the first time. The various computer algorithms included in the invention for automating behavior analysis based on the behavior definitions ensure accurate and efficient identification of mouse behaviors. In addition, the digital video analysis techniques of the present invention improves analysis of behavior by leading to: (1) decreased variance due to non-disturbed observation of the animal; (2) increased experiment sensitivity due to the greater number of behaviors sampled over a much longer time span than ever before possible; and (3) the potential to be applied to all common

normative behavior patterns, capability to assess subtle behavioral states, and detection of changes of behavior patterns in addition to individual behaviors.

The entire behavioral repertoire of individual mice in their home cage was categorized using successive iterations by manual videotape analysis. These
5 manually defined behavior categories constituted the basis of automatic classification. Classification criteria (based on features extracted from the foreground object such as shape, position, movement) were derived and fitted into a decision tree (DT) classification algorithm. The decision tree could classify almost 7000 sample features into 8 different postures classes with accuracy over 94%. A set of HMMs have been
10 built and used to classify the classified postures identified by the DT and yields an almost perfect mapping from input posture to output behaviors in mouse behavior sequences.

The invention may identify some abnormal behavior by using video image information (for example, stored in memory) of known abnormal animals to build a
15 video profile for that behavior. For example, video image of vertical spinning while hanging from the cage top was stored to memory and used to automatically identify such activity in mice. Further, abnormalities may also result from an increase in any particular type of normal behavior. Detection of such new abnormal behaviors may be achieved by the present invention detecting, for example, segments of behavior
20 that do not fit the standard profile. The standard profile may be developed for a particular strain of mouse whereas detection of abnormal *amounts* of a normal behavior can be detected by comparison to the statistical properties of the standard profile.

Thus, the automated analysis of the present invention may be used to build profiles of the behaviors, their amount, duration, and daily cycle for each animal, for example each commonly used strain of mice. A plurality of such profiles may be stored in, for example, a database in a data memory of the computer. One or more of
5 these profiles may then be compared to a mouse in question and difference from the profile expressed quantitatively.

The techniques developed with the present invention for automation of the categorization and quantification of all home-cage mouse behaviors throughout the daily cycle is a powerful tool for detecting phenotypic effects of gene manipulations
10 in mice. As previously discussed, this technology is extendable to other behavior studies of animals and humans, as well as surveillance purposes. As will be described in detail below, the present invention provides automated systems and methods for automated accurate identification, tracking and behavior categorization of an object whose image is captured with video.

15 Other variations of the present invention is directed particularly to automatically determining the behavioral characteristics of an animal in various behavioral experiment apparatus such as water maze, Y-maze, T-maze, zero maze, elevated plus maze, locomotion open field, field for object recognition study, and cued or conditioned fear. In these experiment apparatuses, animal's body contour, center of mass, body parts including head, tail, forelimbs, hind limbs and etc. are
20 accurately identified using the embodiments above. This allows excellent understanding of animal's behaviors within these specific experiment apparatus and procedures. Many novel and important parameters, which were beyond reach previously, are now successfully analyzed. These parameters include, but not limited

to, traces of path of animal's center of mass, instant and average speed, instant and average of body turning angles, distance traveled, turning ratio, proximity score, heading error, stretch-and-attend, head-dipping, stay-across-arms, supported-rearing, sniffing (exploring) at particular objects, latency time to get to the goal (platform),
5 time spent in specific arm/arena or specific zones within arm/arena, number of time entering and exiting arm/arena or specific zones within arm/arena, and etc. These parameters provide good indications for gene targeting, drug screening, toxicology research, learning and memory process study, anxiety study, understanding and treatment of diseases such as Parkinson's Diseases, Alzheimer Disease, ALS, and etc.

10

BRIEF DESCRIPTION OF THE DRAWINGS

Figure 1 is a block diagram of one exemplary system configurable to find the position, shape, and behavioral characteristics of an object using automated video analysis, according to one embodiment of the present invention.

Figure 2 is a block diagram of various functional portions of a computer
15 system, such as the computer system shown in Figure 1, when configured to find the position, shape, and behavioral characteristics of an object using automated video analysis, according to one embodiment of the present invention.

Figure 3 is a flow chart of a method of automatic video analysis for object identification and characterization, according to one embodiment of the present
20 invention.

Figure 4 is a flow chart of a method of automatic video analysis for object identification and characterization, according to another embodiment of the present invention.

Figure 5 is a flow chart of a method of automatic video analysis for object detection and identification, according to one variation of the present invention.

Figure 6 illustrates a sample video image frame with a mouse in a rearing up posture as determined using one variation of the present invention to monitor and
5 characterize mouse behavior.

Figure 7B is a difference image between foreground and background for the image shown in Figure 7A, according to one variation of the present invention as applied for monitoring and characterizing mouse behavior.

Figure 7C is the image shown in Fig. 7A after completing a threshold process
10 for identifying the foreground image of the mouse which is shown as correctly identified, according to one variation of the present invention as applied for monitoring and characterizing mouse behavior.

Figure 7D is a computer generated image showing the outline of the foreground mouse shown in Figure 7A after edge segmentation to demonstrate a
15 contour-based approach to object location and outline identification, according to one variation of the present invention as applied for monitoring and characterizing mouse behavior.

Figure 8 is a chart illustrating one example of various mouse state transitions used in characterizing mouse behavior including: Horizontal Side View Posture (HS);
20 Cuddled Up Posture (CU); Partially Reared Posture (PR); Rear Up Posture (RU); and Horizontal Front/Back View Posture (FB), along with an indication of duration of these states based on a sample, according to one variation of the present invention as applied for monitoring and characterizing mouse behavior.

Figure 9 shows the contour segmentation approach. The contour outline of the animal is split in smaller segments and each segment is classified as a body part.

Figure 10 shows another embodiment in night light conditions. Night conditions are simulated using dim red light.

5 Figure 11 shows another embodiment of the invention, a high-throughput system. Multiple cages can be analyzed at the same time.

DESCRIPTION OF THE PREFERRED EMBODIMENTS

The past few years have seen an increase in the integration of video camera and computer technologies. Today, the integration of the two technologies allows
10 video images to be digitized, stored, and viewed on small inexpensive computers, for example, a personal computer. Further, the processing and storage capabilities of these small inexpensive computers has expanded rapidly and reduced the cost for performing data and computational intensive applications. Thus, video analysis systems may now be configured to provide robust surveillance systems that can
15 provide automated analysis and identification of various objects and characterization of their behavior. The present invention provides such systems and related methods.

In general, the present invention can automatically find the patterns of behaviors and/or activities of a predetermined object being monitored using video. The invention includes a system with a video camera connected to a computer in
20 which the computer is configured to automatically provide object identification, object motion tracking (for moving objects), object shape and posture classification, and behavior identification. In a preferred embodiment the system includes various video analysis algorithms. The computer processes analyze digitized video with the various algorithms so as to automatically monitor a video image to identify, track and

classify the actions of one or more predetermined objects and its movements captured by the video image as it occurs from one video frame or scene to another. The system may characterize behavior by accessing a database of object information of known behavior of the predetermined object. The image to be analyzed may be provided in
5 real time from one or more camera and/or from storage.

In various exemplary embodiments described in detail as follows, the invention is configured to enable monitoring and classifying of animal behavior that result from testing drugs and genetic mutations on animals. However, as indicated above, the system may be similarly configured for use in any of a number of
10 surveillance or other applications. For example, the invention can be applied to various situations in which tracking moving objects is needed. One such situation is security surveillance in public areas like airports, military bases, or home security systems. The system may be useful in automatically identifying and notifying proper law enforcement officials if a crime is being committed and/or a particular behavior
15 being monitored is identified. The system may be useful for monitoring of parking security or moving traffic at intersections so as to automatically identify and track vehicle activity. The system may be configured to automatically determine if a vehicle is speeding or has performed some other traffic violation. Further, the system may be configured to automatically identify and characterize human behavior
20 involving guns or human activity related to robberies or thefts. Similarly, the invention may be capable of identifying and understanding subtle behaviors involving portions of body such as forelimb and can be applied to identify and understand human gesture recognition. This could help deaf individuals communicate. The invention may also be the basis for computer understanding of human gesture to

enhance the present human-computer interface experience, where gestures will be used to interface with computers. The economic potential of applications in computer-human interface applications and in surveillance and monitoring applications is enormous.

5 In one preferred embodiment illustrated in Figure 1, the invention includes a system in which an analog video camera 105 and a video storage/retrieval unit 110 may be coupled to each other and to a video digitization/compression unit 115. The video camera 105 may provide a real time video image containing an object to be identified. The video storage/retrieval unit 110 may be, for example, a VCR, DVD,
10 CD or hard disk unit. The video digitization/compression unit 115 is coupled to a computer 150 that is configured to automatically monitor a video image to identify, track and classify the actions (or state) of the object and its movements (or stillness) over time within a sequence of images. The digitization/compression unit 115 may convert analog video and audio into, for example, MPEG format, Real Player format,
15 etc. The computer may be, for example, a personal computer, using either a Windows platform or a Unix platform, or a MacIntosh computer and compatible platform. In one variation the computer may include a number of components such as (1) a data memory 151, for example, a hard drive or other type of volatile or non-volatile memory; (2) a program memory 152, for example, RAM, ROM, EEPROM,
20 etc. that may be volatile or non-volatile memory; (3) a processor 153, for example, a microprocessor; and (4) a second processor to manage the computation intensive features of the system, for example, a math coprocessor 154. The computer may also include a video processor such as an MPEG encoder/decoder. Although the computer 150 has been shown in Figure 1 to include two memories (data memory 151 and

program memory 152) and two processors (processor 153 and math co-processor 154), in one variation the computer may include only a single processor and single memory device or more than two processors and more than two memory devices. Further, the computer 150 may be equipped with user interface components such as a
5 keyboard 155, electronic mouse 156, and display unit 157.

In one variation, the system may be simplified by using all digital components such as a digital video camera and a digital video storage/retrieval unit 110, which may be one integral unit. In this case, the video digitization/compression unit 115 may not be needed.

10 The computer is loaded and configured with custom software program(s) (or equipped with firmware) using, for example, MATLAB or C/C++ programming language, so as to analyze the digitized video for object identification and segmentation, tracking, and/or behavior/activity characterization. This software may be stored in, for example, a program memory 152 or data memory that may include
15 ROM, RAM, CD ROM and/or a hard drive, etc. In one variation of the invention the software (or firmware) includes a unique background subtraction method which is more simple, efficient, and accurate than those previously known which will be discussed in detail below. In any case, the algorithms may be implemented in software and may be understood as unique functional modules as shown in Figure 2
20 and now described.

Referring to Figure 2, the system is preloaded with standard object information before analyzing an incoming video including a predetermined object, for example, a mouse. First, a stream of digital video including a known object with known characteristics may be fed into the system to a standard object classifier

module 220. A user may then view the standard object on a screen and identify and classify various behaviors of the standard object, for example, standing, sitting, lying, normal, abnormal, etc. Data information representing such standard behavior may then be stored in the standard object behavior storage modules 225, for example a
5 database in data memory 151. Of course, standard object behavior information data sets may be loaded directly into the standard object behavior storage module 225 from another system or source as long as the data is compatible with the present invention protocols and data structure. In any case, once the standard object behavior data is entered into the standard object behavior storage module 225, the system may
10 be used to analyze and classify the behavior of one or more predetermined objects, for example, a mouse.

In the automatic video analysis mode, digital video (either real-time and/or stored) of monitored objects to be identified and characterized is input to an object identification and segregation module 205. This module identifies and segregates a
15 predetermined type of object from the digital video image and inputs it to an object tracking module 210. The object tracking module 210 facilitates tracking of the predetermined object from one frame or scene to another as feature information. This feature information is then extracted and input to the object shape and posture classifier 215. This module classifies the various observed states of the
20 predetermined object of interest into various shape and posture categories and sends it to the behavior identification module 230. The behavior identification module 230 compares the object shape, motion, and posture information with shape, motion, and posture information for a standard object and classifies the behavior accordingly into the predefined categories exhibited by the standard object, including whether the

behavior is normal, abnormal, new, etc. This information is output to the user as characterized behavior information on, for example, a display unit 157.

Referring now to Figure 3, a general method of operation for one embodiment of the invention will be described. In operation, in the video analysis mode the system may receive incoming video images at step 305, from the video camera 105 in real time, pre-recorded from the video storage/retrieval unit 110, and/or a memory integral to the computer 150. If the video is in analog format, then the information is converted from analog to digital format and may be compressed by the video digitization/compression unit 115. The digital video images are then provided to the computer 150 for various computational intensive processing to identify and segment a predetermined object from the image. In a preferred embodiment, the object to be identified and whose activities are to be characterized is a moving object, for example a mouse, which has some movement from frame to frame or scene to scene in the video images and is generally in the foreground of the video images. In any case, at step 310 the digital images may be processed to identify and segregate a desired (predetermined) object from the various frames of incoming video. This process may be achieved using, for example, background subtraction, mixture modeling, robust estimation, and/or other processes.

Next, at step 315, various movements (or still shapes) of the desired object may then be tracked from one frame or scene to another frame or scene of video images. As will be discussed in more detail below, this tracking may be achieved by, for example, tracking the outline contour of the object from one frame or scene to another as it varies from shape to shape and/or location to location. Next, at step 320, the changes in the motion of the object, such as the shapes, locations, and postures of

the object of interest, may be identified and their features extracted and classified into meaningful categories. These categories may include, for example, vertical positioned side view, horizontal positioned side view, vertical positioned front view, horizontal positioned front view, moving left to right, etc. Then, at step 325, the states of the object, for example the shape, location, and posture categories, may be used to characterize the objects activity into one of a number of pre-defined behaviors. For example, if the object is an animal, some pre-defined normal behaviors may include sleeping, eating, drinking, walking, running, etc., and pre-defined abnormal behavior may include spinning vertical, jumping in the same spot, etc. The pre-defined behaviors may be stored in a database in the data memory 151.

Types of behavior may also be characterized using, for example, approaches such as rule-based label analysis, token parsing procedure, and/or Hidden Markov Modeling (HMM). The HMM is particularly helpful in characterizing behavior that is determined with temporal relationships of the various motion of the object across a selection of frames. From these methods, the system may be capable of characterizing the object behavior as new behavior and particular temporal rhythm.

Referring now to Figure 4 a more detailed description of another preferred embodiment will be described. In this case the system is directed toward video analysis of animated objects such as animals. As a preliminary matter, at step 415 video of the activities of a standard object and known behavior characteristics are input into the system. This information may be provided from a video storage/retrieval unit 110 in digitized video form into a standard object classified module 220. This information may then be manually categorized at step 416 to define normal and abnormal activities or behaviors by a user viewing the video

images on the display unit 157 and inputting their classifications. For example, experts in the field may sit together watching recorded scenes. They may then define, for example, an animal's (e.g., a mouse) behavior(s), both qualitatively and quantitatively, with or without some help from systems like the Noldus Observer
5 system. These cataloged behaviors may constitute the important posture and behavior database and are entered into a storage, for example a memory, of known activity of the standard object at step 420. This information provides a point of reference for video analysis to characterize the behavior of non-standard objects whose behaviors/activities need to be characterized such as genetically altered or drug
10 administered mice. For example, normal postures and behaviors of the animals are defined and may be entered into a normal postures and behaviors database.

Once information related to characterizing a standard object(s) is established, the system may then be used to analyze incoming video images that may contain an object for which automated behavior characterization is desired. First, at step 405,
15 incoming video images are received. Next, at decision step 406, the system determines if the video images are in analog or digital format. If the video images are in analog format they are then digitized at step 407. The video may be digitized and may be compressed, using, for example, a digitizer/compression unit 115 into a convenient digital video format such as MPEG, RealPlayer, etc. Otherwise, the
20 digital video image may be input directly to the computer 150. Now the object of interest is identified within the video images and segregated for analysis. As such, at step 408, a background may be generated or updated from the digital video images and foreground objects including a predetermined object for behavior characterization may be detected. For example, a mouse in a cage is detected in the foreground and

segregated from the background. Then, at step 409, features such as centroid, the principal orientation angle of the object, the area (number of pixels), the eccentricity (roundness), and the aspect ratio of the object, and/or shape in terms of contour, convex hull, or b-spline, of the foreground object of interest (e.g., a mouse) are
5 extracted. Next, at step 410, the foreground object shape and postures are classified into various categories, for example, standing, sitting, etc.

Then, at step 411, the foreground object (e.g., a mouse) posture may be compare to the various predefined postures in the set of known postures in the standard object storage of step 420, which may be included in a database. At steps
10 412, the observed postures of the object contained in the analyzed video image may be classified and identified as a particular posture known for the standard object or a new previously unidentified posture. Next, at step 413, various groups of postures may be concatenated into a series to make up a foreground object behavior that is then compared against the sequence of postures, stored in for example a database in
15 memory, that make up a known standard object behavior. This known standard behavior is, in a preferred embodiment, normal behavior for the type of animal being studied. However, the known activity of the standard object may be normal or abnormal behavior of the animal. In either case, at step 414, the abnormal behaviors are then identified in terms of (1) known abnormal behavior; (2) new behavior likely
20 to be abnormal; and/or (3) daily rhythm differences likely to be abnormal behavior. Known normal behavior may also be output as desired by the user. This information is automatically identified to the user for their review and disposition. In one variation of the invention, the information output may include behavior information that is compatible with current statistical packages such as Systat and SPSS.

In one embodiment of the invention as illustrated in Figure 5, object detection is performed through a unique method of background subtraction. First, at step 405, incoming video is provided to the system for analysis. This video may be provided by digital equipment and input to the object identification and segregation module 5 205 of the computer 150. Next, at step 505, the incoming digital video signal may be split into individual images (frames) in real-time. This step may be included if it is desired to carry out real-time analysis. Then, at decision step 506, the system determines if the background image needs to be developed because there was no background image developed previously or the background image has changed. If 10 the background image needs to be generated or updated, then at step 507 a background image is generated by first grouping a number of frames or images into a sample of video images, for example 20 frames or images. The background may need to be updated periodically due to changes caused by, for example, lighting and displacement of moveable objects in the cage, such as the bedding. Then, at step 508 15 the system generates a standard deviation map of the group of images. Next, at step 509, an object(s) bounding box area is identified and removed from each frame or image to create a modified frame or image. The bounding box area is determined by sensing the area wherein the variation of a feature such as the standard deviation of intensity is above a predetermined threshold. Thus, an area in the digitized video 20 image where the object of interest in motion is located is removed leaving only a partial image. Then, at step 510, the various modified images within the group, less the bounding box area, are combined, for example averaged, to create a background image at step 511.

Since varying pixels are not used in averaging, "holes" will be created in each image that is being used in the averaging process. Over time, not all frames will have these holes at the same location and hence, a complete background image is obtained after the averaging process. Final background is obtained by averaging 5-10 samples.

5 This completes at least one iteration of the background generation process.

The background image does not remain constant for a great length of time due to various reasons. For example, the bedding in a mouse cage can shift due to the activity of the mouse. External factors such as change in illumination conditions also require background image recalculations. If the camera moves, then, background
10 might need to be changed. Thus, the background typically needs to be recalculated periodically as described above or it can be recalculated by keeping track of the difference image and note any sudden changes such as an increase in the number of particular color (e.g., white) pixels in the difference image or the appearance of patches of the particular color (e.g., white) pixels in another area of the difference
15 image. In any case, the newly generated background image may then be combined with any existing background image to create a new background image at step 511.

The newly generated background image is next, at step 512, subtracted from the current video image(s) to obtain foreground areas that may include the object of interest. Further, if the background does not need to be updated as determined at
20 decision step 506, then the process may proceed to step 512 and the background image is subtracted from the current image, leaving the foreground objects.

Next, at steps 513-518, the object identification/detection process is performed. First, at step 513, regions of interest (ROI) are obtained by identifying an area where the intensity difference is greater than a predetermined threshold, which

constitute potential foreground object(s) being sought. Classification of these foreground regions of interest will be performed using the sizes of the ROIs, distances among these ROIs, threshold of intensity, and connectedness to identify the foreground objects. Next, the foreground object identification/detection process may
5 be refined by utilizing information about the actual distribution (histograms) of the intensity levels of the foreground object and using edge detection to more accurately identify the desired object(s).

At step 514, during both the background generation and background subtraction steps for object identification, the system continuously maintains a
10 distribution of the foreground object intensities as obtained. A lower threshold may be used to thereby permit a larger amount of noise to appear in the foreground image in the form of ROIs. Thus, at step 514, a histogram is then updated with the pixels in the ROI. At step 515, plotting a histogram of all the intensities of a particular color pixels over many images, provides a bi-modal shape with the larger peak
15 corresponding to the foreground object's intensity range and the smaller peak corresponding to the noise pixels in the ROI's images. Now, at step 516, having "learned" the intensity range of the foreground object, only those pixels in the foreground object that conform to this intensity range are selected, thereby identifying the foreground object more clearly even with background that is fairly similar.

20 In any case, next at step 517 the foreground object of interest may be refined using edge information to more accurately identify the desired object. An edge detection mechanism such as Prewitt operator is applied to the original image. Adaptive thresholds for edge detections can be used. Once the edge map is obtained, the actual boundary of the foreground object is assumed to be made up of one or more

segments in the edge map, i.e., the actual contour of the foreground objects comprises edges in the edge map. The closed contour of the "detected" foreground object is broken into smaller segments, if necessary. Segments in the edge map that are closest to these contour segments according to a distance metric are found to be the desired
5 contour. One exemplary distance metric is the sum of absolute normal distance to the edge map segment from each point in the closed contour of the "detected" foreground object. Finally, at step 518 the information identifying the desired foreground object is output. The process may then continue with tracking and/or behavior characterization steps.

10 The previous embodiments are generally applicable to identifying, tracking, and characterizing the activities of a particular object of interest present in a video image, e.g., an animal, a human, a vehicle, etc. However, the invention is also particularly applicable to the study and analysis of animals used for testing new drugs and/or genetic mutations. As such, a number of variations of the invention related to
15 determining changes in behavior of mice will be described in more detail below using examples of video images obtained.

One variation of the present invention is designed particularly for the purpose of automatically determining the behavioral characteristics of a mouse. The need for sensitive detection of novel phenotypes of genetically manipulated or drug-
20 administered mice demands automation of analyses. Behavioral phenotypes are often best detected when mice are unconstrained by experimenter manipulation. Thus, automation of analysis of behavior in a home cage would be a preferred means of detecting phenotypes resulting from gene manipulations or drug administrations. Automation of analysis as provided by the present invention will allow quantification

of all behaviors and may provide analysis of the mouse's behavior as they vary across the daily cycle of activity. Because gene defects causing developmental disorders in humans usually result in changes in the daily rhythm of behavior, analysis of organized patterns of behavior across the day may be effective in detecting
5 phenotypes in transgenic and targeted mutant mice. The automated system of the present invention may also detect behaviors that do not normally occur and present the investigator with video clips of such behavior without the investigator having to view an entire day or long period of mouse activity to manually identify the desired behavior.

10 The systematically developed definition of mouse behavior that is detectable by the automated analysis of the present invention makes precise and quantitative analysis of the entire mouse behavior repertoire possible for the first time. The various computer algorithms included in the invention for automating behavior analysis based on the behavior definitions ensure accurate and efficient identification
15 of mouse behaviors. In addition, the digital video analysis techniques of the present invention improves analysis of behavior by leading to: (1) decreased variance due to non-disturbed observation of the animal; (2) increased experiment sensitivity due to the greater number of behaviors sampled over a much longer time span than ever before possible; and (3) the potential to be applied to all common normative behavior
20 patterns, capability to assess subtle behavioral states, and detection of changes of behavior patterns in addition to individual behaviors. Development activities have been complete to validate various scientific definition of mouse behaviors and to create novel digital video processing algorithms for mouse tracking and behavior

recognition, which are embodied in software and hardware system according to the present invention.

Various lighting options for videotaping have been evaluated. Lighting at night as well as with night vision cameras was evaluated. It has been determined that good quality video was obtained with normal commercial video cameras using dim red light, a frequency that is not visible to rodents. Videos were taken in a standard laboratory environment using commercially available cameras 105, for example a Sony analog camera, to ensure that the computer algorithms developed would be applicable to the quality of video available in the average laboratory. The commercially available cameras with white lighting gave good results during the daytime and dim red lighting gave good results at night time.

Referring again to Figure 3, the first step in the analysis of home cage behavior is an automated initialization step that involves analysis of video images to identify the location and outline of the mouse, as indicated by step 310. Second, the location and outline of the mouse are tracked over time, as indicated by step 315. Performing the initialization step periodically may be used to reset any propagation errors that appear during the tracking step. As the mouse is tracked over time, its features including shape are extracted, and used for training and classifying the posture of the mouse from frame to frame, as indicated by step 320. Posture labels are generated for each frame, which are analyzed over time to determine the actual behavior, as indicated by step 325. The steps 305, 310, and 315 have been presented in the earlier application, and hence it will only be described very briefly. The steps 320 and 325 will then be described in detail using the particular application of mouse behavior characterization. Detailed descriptions of how each of the behaviors is

modeled, and the corresponding methodology of detecting each of the behaviors in the repertoire are presented before step 325.

5 I. Location and Outline Identification and Feature Extraction

The first step in analyzing a video of an animal and to analyze the behavior of the animal is to locate and extract the animal. A pre-generated background of the video clip in question is first obtained and it is used to determine the foreground objects by taking the intensity difference and applying a threshold procedure to
10 remove noise. This step may involve threshold procedures on both the intensity and the size of region. An 8-connection labeling procedure may be performed to screen out disconnected small noisy regions and improve the region that corresponds to the mouse. In the labeling process, all pixels in a frame will be assigned a label as foreground pixel or background pixel based on the threshold. The foreground pixels
15 are further cleaned up by removing smaller components and leaving only the largest component as the foreground object. Those foreground pixels that border a background pixel form the contour for the object. The outline or contour of this foreground object is thus determined. The centroid (or center of mass) of the foreground object is calculated and is used for representing the location of the object
20 (e.g., mouse).

Figures 7A, 7B, 7C, and 7D illustrate the results of the location and object outline identification for a mouse using the present invention. Figure 7B illustrates a difference image between foreground and background for the image in Figure 7A. Figure 7C illustrates the image after thresholding showing the foreground mouse
705 object correctly identified. Figure 7D illustrates the extracted contour of this object.
25

The contour representation can be used as features of the foreground object, in addition to other features that include but not limited to: centroid, the principal orientation angle of the object, the area (number of pixels), the eccentricity (roundness), and the aspect ratio of object.

5

II. Mouse tracking

Ideal tracking of foreground objects in the image domain involves a matching operation to be performed that identifies corresponding points from one frame to the next. This process may become computationally too consuming or expensive to perform in an efficient manner. Thus, one approach is to use approximations to the ideal case that can be accomplished in a short amount of time. For example, tracking the foreground object may be achieved by merely tracking the outline contour from one frame to the next in the feature space (i.e., identified foreground object image).

In one variation of the invention, tracking is performed in the feature space, which provides a close approximation to tracking in the image domain. The features include the centroid, principal orientation angle of the object, area (number of pixels), eccentricity (roundness), and the aspect ratio of object with lengths measured along the secondary and primary axes of the object. In this case, let S be the set of pixels in the foreground object, A denote the area in number of pixels, (C_x, C_y) denote the centroid, ϕ denote the orientation angle, E denote the eccentricity, and R denote the aspect ratio. Then,

20

$$C_x = \frac{1}{A} \sum_S x \quad C_y = \frac{1}{A} \sum_S y$$

Let us define three intermediate terms, called second order moments,

$$m_{2,0} = \sum_S (x - C_x)^2 \quad m_{0,2} = \sum_S (y - C_y)^2 \quad m_{1,1} = \sum_S (x - C_x)(y - C_y)$$

Using the central moments, we define,

$$\phi = \frac{1}{2} \arctan \frac{2m_{1,1}}{m_{2,0} - m_{0,2}}$$

5

$$E = \frac{(m_{2,0} - m_{0,2})^2 + 4m_{1,1}^2}{(m_{2,0} + m_{0,2})^2}$$

R is equal to the ratio of the length of the range of the points projected along an axis perpendicular to ϕ , to the length of the range of the points projected along an axis parallel to ϕ . This may also be defined as the aspect ratio (ratio of width to length) after rotating the foreground object by ϕ .

10

Tracking in the feature space involves following feature values from one frame to the next. For example, if the area steadily increases, it could mean that the mouse is coming out of a cuddled up position to a more elongated position, or that it could be moving from a front view to a side view, etc. If the position of the centroid of the mouse moves up, it means that the mouse may be rearing up on its hind legs.

15

Similarly, if the angle of orientation changes from horizontal to vertical, it may be rearing up. These changes can be analyzed with combinations of features also.

However, it is possible for a contour representation to be used to perform near-optimal tracking efficiently in the image domain (i.e., the complete image before background is subtracted).

20

III. Mouse posture classification

Once the features are obtained for the frames in the video sequence, the foreground state of the mouse is classified into one of the given classes. This involves building a classifier that can classify the shape using the available features.

5 This information may be stored in, for example, a database in, for example, a data memory. In one variation of the invention a Decision Tree classifier (e.g., object shape and posture classifier 215) was implemented by training the classifier with 6839 samples of digitized video of a standard, in this case, normal mouse. Six attributes (or features) for each sample were identified. Ten posture classes for
10 classification were identified as listed below.

1. Horizontal Side View Posture – Horizontally positioned, side view, either in normal state or elongated.
2. Vertical Posture – Vertically positioned in a reared state (e.g., See Figure 6).
- 15 3. Cuddled Posture – Cuddled up position (like a ball).
4. Horizontal Front/Back View Posture – Horizontally positioned, but either front or back view, i.e., axis of mouse along the viewer's line of sight.
5. Partially Reared Posture – Partially reared, e.g., when drinking or eating, sitting on hind legs (e.g., See Figure 7A).
- 20 6. Stretched Posture – Stretched horizontally or vertically.
7. Hang Vertical Posture – Hanging vertically from the top of the cage or food bin.
8. Hang Cuddled Posture – Hanging cuddled up close to the top of the cage or on the food bin.
- 25 9. Eating Posture – In one of the earlier 8 posture with the added condition that the mouth is in touch with the food bin.
10. Drinking Posture – In one of the postures 1-8 with the added condition that the mouth is in touch with the water spout.

30 The system of the present invention was exercised using these classifications.

Performing a 10-fold cross-validation on the 6839 training samples, a combined accuracy of 94.6% was obtained indicating that the classifier was performing well.

This is in the range of the highest levels of agreement between human observers. The

present system provides good accuracy for mouse shape and posture recognition and classification.

After the posture is classified, various body parts of the animal that can be obtained from that posture is detected. The contour of the animal object is split into smaller segments based on the curvature features. Segments are split at concave points along the contour. A segment comprising those contour pixels starting from an extreme concave point to the next extreme concave point and containing an extreme convex point is considered as a body segment. These body segments are classified into one of the following classes: Head, Forelimb, Abdomen, Hind Limb, Tail, Lower Back, Upper Back, and Ear.

With the combination of the posture information and the body part information from a plurality of frames, behaviors are modeled and detected.

IV. Behavior Detection Methodology

A typical video frame of a mouse in its home cage is shown in Figure 6. In this video frame a mouse is shown in a rearing up posture. Many such frames make up the video of, for example, a 24 hour mouse behavior monitoring session. A small segment of successive frames of this video will correspond to one of the behaviors in the group of behaviors that have been modeled. The approach is to identify the correct segments and how to match those segments to the correct behavior. How each behavior is modeled is first described.

Each behavior can be modeled as a sequence of postures of the mouse. If this particular pattern of postures is exhibited by the mouse, the corresponding behavior is

detected. The following set of postures is being used: Horizontal Side View Posture, Vertical Posture, Cuddled Posture, Horizontal Front/Back View Posture, Partially Reared Posture, Stretched Posture, Hang Vertical Posture, Hang Cuddled Posture, Eating Posture and Drinking Posture. Apart from modeling a behavior as a sequence

5 of postures, certain rules or conditions can be attached to the behavior description, which, only if satisfied will determine the corresponding behavior. The rules or conditions can be formulated using any of the available features or parameters including position and shape of specific body parts with or without respect to other objects, motion characteristics of the entire mouse body or individual body parts, etc.

10 In the descriptions below, all such rules or conditions that augment the posture sequence requirement to derive the specific modeling of the behavior are stated. The behavior descriptions follow:

A. Rear Up to a Full or a Partially Reared Posture

15 Rear Up behavior is modeled as a sequence of postures starting from either of the cuddled, horizontal side-view, or horizontal front/back view postures to ending in a vertical or partially reared posture. This behavior is analogous to the standing up behavior.

20 B. Come Down Fully or to a Partially Reared Posture

Come Down behavior is modeled as a sequence of postures starting from either vertical or partially reared posture to ending in one of cuddled, horizontal side view or horizontal front/back view postures. This behavior is analogous to the sitting down or laying down behavior.

C. Eat

Eating behavior is modeled as a sequence of eating postures. An eating posture is an augmentation of one of the other postures by a condition that the mouth body part of
5 the mouse is in contact with a food access area in the cage.

D. Drink

Drinking behavior is modeled as a sequence of drinking postures. A drinking posture is an augmentation of one of the other postures by a condition that the mouth body
10 part of the mouse is contact with a water spout in the cage.

E. Dig

Digging behavior is determined by the aft movement of the bedding material in the cage by the animal with its fore and hind limbs. The displacement of the bedding is
15 detected and the direction of movement of the bedding along with the orientation of the mouse is used to determine this behavior.

F. Forage

Foraging behavior is determined by the movement of bedding material in the cage by
20 the animal using the head and forelimbs. The displacement of the bedding is detected along with the position of the head and forelimbs and this is used to determine the foraging behavior.

G. Jump

Jump behavior is modeled by a single up and down movement of the animal. Both the top of the animal and the bottom of the animal have to move monotonously up, and then, down, to determine this behavior.

5 H. Jump Repetitively

Repetitive jumping behavior is determined by several continuous up and down movements (individual jumps) of the animal.

I. Sniff

10 Sniffing behavior is determined by a random brisk movement of the mouth/nose tip of the head while the rest of the body remains stationary. The trace of the mouth tip is analyzed and the variance in its position is high relative to the bottom of the animal, a sniff is detected.

15 J. Hang

Hang behavior is modeled as a sequence of postures starting from the vertical posture to ending in a hang vertical or hang cuddled posture.

K. Land after Hanging

20 Land behavior is modeled as a sequence of postures starting from the hang vertical or hang cuddled posture to ending in a vertical posture.

L. Sleep

25 Sleep behavior is detected by analyzing the contour of the mouse body. If the amount of movement of this contour from one frame to the next is below a threshold value for a prolonged period of time, the mouse enters a sleep state.

M. Twitch during Sleep

Twitch behavior is determined by the detection of a brief period of substantial movement and the resumption of sleep activity following this brief movement.

5

N. Awaken from Sleep

Awaken behavior is determined by a prolonged substantial movement of the animal after sleep had set in.

10

O. Groom

Grooming behavior is modeled as a brisk movement of limbs and head in a cyclical and periodic pattern. Variances of several shape and motion parameters, including the width and height, and area of the mouse, are calculated over time and if these variances exceed a threshold, for a prolonged period of time, groom is detected.

15

P. Pause briefly

Pause behavior is determined by a brief absence of movement of the animal. Similar criteria as those used for sleep detection is employed, except the duration of the behavior is much shorter, only lasting for several seconds.

20

Q. Urinate

Urinate behavior is determined by the detection of the mouse tail being raised up and the mouse remaining stationary briefly while the tail is up.

25

R. Turn

Turn behavior is modeled as a sequence of postures starting from horizontal side view or cuddled posture to ending in a horizontal front/back view posture, or vice versa.

Accordingly, the turn behavior can further be classified as a Turn to Face Right, Turn to Face Left, Turn to Face Forward or Back behavior.

S. Circle

5 Circling behavior is modeled as a succession of 3 or more turns.

T. Walk

Walking or running behavior is determined by the continuous sideways movement of the centroid of the animal in one direction, to the left or right. The mouse needs to
10 travel a certain minimal distance over a specified length of time for this behavior to be detected.

U. Stretch

Stretch behavior is modeled as a sequence of Stretched Postures. A Stretched posture
15 is determined by the observation of the upper and lower back contours. If for a given frame, those body parts have a concave shape instead of a normal convex shape, and the overall shape of the animal is elongated, then a Stretched posture is detected for that frame. A sequence of these Stretched postures generates a Stretch behavior. This Stretch behavior can occur when the animal is horizontally elongated or
20 vertically elongated. Horizontally elongated Stretching behavior occur after awaken behavior or when ducking under objects. Vertically elongated Stretch behavior occurs during sniffs or supported rearing behaviors.

V. Chew

Chewing behavior is modeled as a movement of the mouth while the mouth is not in touch with a food container. Chews are detected only between two co-occurring Eat behaviors.

5 W. Stationary

Stationary behavior is detected when the animal remains in the same place and does not perform any of the other behaviors. It is often output as a default behavior when no other behavior can be detected. But, if the mouse moves and the movement pattern does not match any of the other behaviors, Unknown Behavior, not Stationary
10 behavior, is selected.

 X. Unknown Behavior

If the activity cannot be characterized by any of the behavior models, the behavior is deemed to be unknown

15 V. Behavior identification

Using the posture labels assigned for the frames in the video clip, the approach is to determine those pre-defined behaviors as defined in the previous step. This process will be accomplished in real-time so that immediate results will be
20 reported to investigators or stored in a database. One approach is to use a rule-based label analysis procedure (or a token parsing procedure) by which the sequence of labels is analyzed and to identify particular behaviors when its corresponding sequence of labels is derived from a video frame being analyzed. For example, if a long sequence (lasting for example several minutes) of the “Cuddled up position”
25 label (Class 3) is observed, and if its centroid remains stationary, then, it may be

concluded that the mouse is sleeping. If the location of the waterspout is identified, and if we observe a series of “partially reared” (Class 5) labels, and if the position of the centroid, and the mouse’s angle of orientation fall within a small range that has been predetermined, the system can determine and identify that the mouse is drinking.

5 It may also be useful for certain extra conditions to be tested such as, “some part (the mouth) of the mouse must touch the spout if drinking is to be identified” in addition to temporal characteristics of the behavior.

While this approach is very straightforward, a better approach involves using a probabilistic model such as Hidden Markov Models (HMMs), where models may
10 be built for each class of behavior with training samples. These models may then be used to identify behaviors based on the incoming sequence of labels. The HMM can provide significant added accuracy to temporal relationships for proper complex behavior characterization.

Referring now to Figure 8, various exemplary mouse state transitions tested in
15 the present invention are illustrated. The five exemplary mouse state transitions include: (1) Horizontal Side View Posture (HS) 805, (2) Horizontal Front/Back Posture (FB) 810 postures, (3) Cuddled Up Posture (CU) 815, (4) Partially Reared Posture (PR) 820, and (5) Reared Up Posture (RU) 825. As illustrated, Figure 8 shows the five posture states and the duration for which a mouse spent in each state in
20 an exemplary sample video clip. One example of a pattern that is understandable and evident from the figure is that the mouse usually passes through the partially reared posture (PR) 820 state to reach the reared up posture (RU) 825 state from the other three ground-level states. The states are defined according to the five posture classes mentioned previously.

Many important features can be derived from this representation, e.g., if the state changes are very frequent, it would imply that the mouse is very active. If the mouse remained in a single ground-level state such as “cuddled-up” (class 3) for an extended period of time, the system may conclude that the mouse is sleeping or resting. The sequence of transitions are also important, e.g., if the mouse rears (class 5 2) from a ground-level state such as “Horizontally positioned” (class 1), it should pass briefly through the partially reared state (class 5). Techniques such as HMMs exploit these types of time-sequence-dependent information for performing classification.

Each of the behaviors described in the previous section that can be modeled as a sequence of postures, was provided with a trained HMM representing that behavior only. Hence, there was a one-to-one correspondence between each HMM and a behavior that it represented. For example, an HMM corresponding to Rear Up From Partially Reared (RUFP) was created to represent the Rear Up behavior from a partially reared state fully to a reared up state. This was done during the training step.

During HMM training, posture sequence from real-video data was extracted that corresponded to various behaviors. Several samples for each behavior were collected. A separate HMM was generated for each of these behaviors that could be represented by a simple sequence of postures. For example, for a Rear Up From Partially Reared (RUFP) behavior, a sample sequence of postures can be 5, 5, 5, 2, 2, 2, 2, where the numbers represent the posture class described earlier. Similarly, another sample can be 5, 5, 2, 2, 2, 2, 2. More complicated behaviors will have more complicated patterns.

Once trained, these HMMs will match best with a sequence of labels that has a pattern similar to those used for training. For example, an input sequence of the

form, 5, 5, 5, 5, 5, 2, 2, 2 will match with the RUFP better than any other HMM. Hence, during analysis, the incoming sequence of labels is grouped and presented to all the HMMs and the winning HMM (or the best matching HMM) is selected as the corresponding behavior for that frame sequence. Continuing this process, all the
5 behaviors that occur in succession are detected and output.

One of the distinct advantages of using the HMM approach is that noise during analysis does not affect the match values much. So, the sequence 5, 5, 5, 7, 2, 2, 2, will still match with the RUFP HMM better than any other HMM.

If certain augmentation rules needed to be applied, they were applied in a rule-
10 based approach during the real-time analysis. For example, to detect grooming behavior, it is required that the variance of the width, height, and other measures be within a pre-set range while the animal has a certain sequence of postures. If both these conditions – the posture-based condition and the feature-based condition – the grooming behavior is detected.

15 Although the above exemplary embodiment is directed to a mouse analyzed in a home cage, it is to be understood that the mouse (or any object) may be analyzed in any location or environment. Further, the invention in one variation may be used to automatically detect and characterize one or more particular behaviors. For example, the system could be configured to automatically detect and characterize an animal
20 freezing and/or touching or sniffing a particular object. Also, the system could be configured to compare the object's behavior against a “norm” for a particular behavioral parameter. Other detailed activities such as skilled reaching and forelimb movements as well as social behavior among groups of animals can also be detected and characterized.

In summary, when a new video clip is analyzed, the system of the present invention first obtains the video image background and uses it to identify the foreground objects. Then, features are extracted from the foreground objects, which are in turn passed to the decision tree classifier for classification and labeling. This
5 labeled sequence is passed to a behavior identification system module that identifies the final set of behaviors for the video clip. The image resolution of the system that has been obtained and the accuracy of identification of the behaviors attempted so far have been very good and resulted in an effective automated video image object recognition and behavior characterization system.

10 The invention may identify some abnormal behavior by using video image information (for example, stored in memory) of known abnormal animals to build a video profile for that behavior. For example, video image of vertical spinning while hanging from the cage top was stored to memory and used to automatically identify such activity in mice. Further, abnormalities may also result from an increase in any
15 particular type of normal behavior. Detection of such new abnormal behaviors may be achieved by the present invention detecting, for example, segments of behavior that do not fit the standard profile. The standard profile may be developed for a particular strain of mouse whereas detection of abnormal amounts of a normal behavior can be detected by comparison to the statistical properties of the standard
20 profile. Thus, the automated analysis of the present invention may be used to build a profile of the behaviors, their amount, duration, and daily cycle for each animal, for example each commonly used strain of mice. A plurality of such profiles may be stored in, for example, a database in a data memory of the computer. One or more of

these profiles may then be compared to a mouse in question and difference from the profile expressed quantitatively.

The techniques developed with the present invention for automation of the categorization and quantification of all home-cage of mouse behaviors throughout the daily cycle is a powerful tool for detecting phenotypic effects of gene manipulations in mice. As previously discussed, this technology is extendable to other behavior studies of animals and humans, as well as surveillance purposes. In any case, the present invention has proven to be a significant achievement in creating an automated system and methods for automated accurate identification, tracking and behavior categorization of an object whose image is captured in a video image.

In another embodiment of the invention, the analysis is performed under simulated night conditions with the use of red-light and regular visible range cameras, or with the use of no-light conditions and infra-red cameras.

In another embodiment of the invention, there are multiple cameras taking video images of experiment cages that contain animals. There is at least one cage, but as many as the computer computing power allows, say four (4) or sixteen (16) or even more, can be analyzed.

The systematically developed definitions of mouse behaviors that are detectable by the automated analysis according to the present invention makes precise and quantitative analysis of the entire mouse behavior repertoire possible for the first time. The various computer algorithms included in the invention for automating behavior analysis based on the behavior definitions ensure accurate and efficient identification of mouse behaviors. In addition, the digital video analysis techniques of the present invention improves analysis of behavior by leading to: (1) decreased

variance due to non-disturbed observation of the animal; (2) increased experiment sensitivity due to the greater number of behaviors sampled over a much longer time span than ever before possible; and (3) the potential to be applied to all common normative behavior patterns, capability to assess subtle behavioral states, and
5 detection of changes of behavior patterns in addition to individual behaviors.

Although particular embodiments of the present invention have been shown and described, it will be understood that it is not intended to limit the invention to the preferred or disclosed embodiments, and it will be obvious to those skilled in the art that various changes and modifications may be made without departing from the spirit
10 and scope of the present invention. Thus, the invention is intended to cover alternatives, modifications, and equivalents, which may be included within the spirit and scope of the invention as defined by the claims.

For example, the present invention may also include audio analysis and/or multiple camera analysis. The video image analysis may be augmented with audio
15 analysis since audio is typically included with most video systems today. As such, audio may be an additional variable used to determine and classify a particular objects behavior. Further, in another variation, the analysis may be expanded to video image analysis of multiple objects, for example mice, and their social interaction with one another. In a still further variation, the system may include
20 multiple cameras providing one or more planes of view of an object to be analyzed. In an even further variation, the camera may be located in remote locations and the video images sent via the Internet for analysis by a server at another site. In fact, the standard object behavior data and/or database may be housed in a remote location and the data files may be downloaded to a stand alone analysis system via the Internet, in

accordance with the present invention. These additional features/functions add versatility to the present invention and may improve the behavior characterization capabilities of the present invention to thereby achieve object behavior categorization which is nearly perfect to that of a human observer for a broad spectrum of applications.

5

All publications, patents, and patent applications cited herein are hereby incorporated by reference in their entirety for all purposes.